NON-PROVISIONAL APPLICATION FOR UNITED STATES PATENT

FOR

Digital Audio Signal Compression Method and Apparatus

<u>Inventor</u>

REZNIK, Yuriy A.

DIGITAL AUDIO SIGNAL COMPRESSION METHOD AND APPARATUS

RELATED APPLICATION

[0001] This is a non-provisional application of provisional application 60/464,068, filed 04/18/03, entitled "Noiseless Compression of PCM Audio Signals". This non-provisional application claims priority to said '068 provisional application.

FIELD OF THE INVENTION

[0002] The present invention relates to the field of signal processing. More specifically, the present invention relates to compression of audio signal data.

BACKGROUND OF THE INVENTION

[0003] Digital audio has a number of advantages over analog audio. In particular, pulse code modulation (PCM) audio has a number of advantages over other audio formats. For example, digital audio, in particular, PCM audio, offers freedom to interchange audio data without generation loss between media. Increasingly, PCM audio is not only being offered from medium like compact disc (CD), it is also widely employed in broadcast programming, through air waves or cable, or in streamed contents, through private and/or public networks, such as the Internet.

[0004] For broadcast programming or streamed contents, bandwidth availability/consumption remains a significant challenge.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] The present invention will be described by way of exemplary embodiments, but not limitations, illustrated in the accompanying drawings in which like references denote similar elements, and in which:

[0006] **Figure 1** illustrates a method view for compressing audio signal data, in accordance with some embodiments of the present invention; and

[0007] **Figure 2** illustrates a system, including its transmit and receive sections, in accordance with some embodiments.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0008] Illustrative embodiments of the present invention include, but are not limited to, method to compress digital audio data (in particular, PCM audio data), encoders/decoders adapted to practice all or portions of the method, and systems having the encoders/decoders.

[0009] In the description to follow, for ease of understanding, the present invention will primarily be described in the context of PCM audio embodiments, however, the present invention may be practiced for other digital audio, e.g. one-bit oversampled audio representations commonly used in super-audio compact disks (SACD).

[0010] Various aspects of embodiments of the present invention will be described. However, various embodiments may be practiced with only some or all of the described aspects. For purposes of explanation, specific numbers, materials and configurations are set forth in order to provide a thorough understanding of the embodiments being described. In alternate embodiments, they may be practiced without the specific details. In various instances, well-known features are omitted or simplified in order not to obscure the essence of the embodiments.

[0011] Parts of the description will be presented in signal processing terms, such as data, filtering, quantization, encoding, decoding, and so forth, consistent with the manner commonly employed by those skilled in the art to convey the substance of their work to others skilled in the art. As well understood by those skilled in the art, the data quantities take the form of electrical, magnetic, or optical signals capable of being stored, transferred, combined, and otherwise manipulated through mechanical, electrical and/or optical components of a general/special purpose computing device/system.

[0012] The term "computing device/system" as used herein includes general purpose as well as special purpose data processing machines, systems, and the like, that are standalone, adjunct or embedded. Examples of general purpose "computing devices/systems" include, but are not limited to, handheld computing devices (palm sized, tablet sized and so forth), laptop computing devices, desktop computing devices, servers, and so forth. Examples of special purpose "computing device/system" include,

but are not limited to, processor based wireless mobile phones, handheld digital music players, set-top boxes, game boxes/consoles, CD/DVD players, digital cameras, digital CAMCORDERs, and so forth. [DVD – Digital Versatile Disk]

[0013] Various operations will be described as multiple discrete operations in turn, in a manner that is most helpful in understanding the various embodiments of the present invention, however, the order of description should not be construed as to imply that these operations are necessarily order dependent. In particular, these operations need not be performed in the order of presentation.

[0014] The phrase "in one embodiment" is used repeatedly. The phrase generally does not refer to the same embodiment; however, it may. The terms "comprising", "having" and "including" are synonymous, unless the context dictates otherwise.

[0015] Referring now to Figure 1, wherein a method view of the present invention, in accordance with some embodiments, is illustrated. As illustrated, for the embodiments, the process 100 starts with the receiving 102 of a portion of a stream of audio signal data (e.g. PCM audio signal data). On receipt, or shortly thereafter, the audio signal data is partitioned 104 into a number of data blocks for subsequent processing (compression). In various embodiments, the audio signal data is partitioned 104 into a number of fixed or variable size data blocks, and when variable sized-blocks are used, the variable data block sizes are conveyed 108 to the recipient (e.g. multiplexed 128 onto the transmission bit stream). For the embodiments, the default fixed data block size is assumed to be known to the recipient, however, in other embodiments, the invention may nonetheless be practiced including the conveyance of the fixed data block size to the recipient (e.g. multiplexed 128 onto the transmission bit stream).

[0016] For the embodiments, the data blocks are selected 106 one block at a time, and the remaining operations of process 100 are applied to the currently selected data block to process and compress the data block, and ultimately, after compression, placing 128 the processed/compressed data onto the transmission bit stream for transmission to a recipient. The operations are repeated until all data blocks of the received portion of the audio signal have been processed (compressed), and multiplexed 128 onto the transmission bit stream.

[0017] In alternate embodiments, each data block may be further partitioned into sub-blocks, with the sub-blocks being selected for processing (compression) and multiplexed **128** onto the transmission bit stream, one sub-block at a time. Likewise, for these embodiments, the sub-block size may be fixed or variable. Regardless, the sub-block size is conveyed to a recipient. For these embodiments, the operations are also repeated until all sub-blocks of the data block have been processed (compressed) and multiplexed **128** onto the transmission bit stream. Then, the operations are repeated again until all data blocks of the received portion of the audio signal have been processed (compressed) and multiplexed **128** onto the transmission bit stream.

[0018] Continuing to refer to **Fig. 1**, on selection, a prediction filter is applied **110** to the unit (block or sub-block)of audio data to be processed and compressed. Similar to the block/sub-block size information, the parameters of the prediction filter are conveyed **112-114** to the recipient (e.g. multiplexed **128** onto the transmission bit stream).

[0019] In various embodiments, the filtering may be assisted by the employment of neighboring blocks. In various embodiments, the prediction filter is a Linear Prediction Filter, and the parameters are the prediction order p, and the prediction coefficients a_1, ... a_p. Inv various embodiments, the parameters conveyed **112-114** to the recipient may include the prediction order p, the quantization step size use to quantize prediction coefficients, and the quantized prediction coefficients $\hat{a}_1,...,\hat{a}_p$.

[0020] As illustrated, for the embodiments, as a result of the application of the prediction filter to the unit of audio data, residual samples e_1, ... e_n are generated **116**. Next, a number of statistical measures are determined for the residual samples to characterize **118** their distribution (to be described more fully below). For the embodiment, the statistical measures are employed to form **120** a distribution descriptor (also to be described more fully below), which in turn is conveyed to the recipient (e.g. multiplexed onto the transmission bit stream). Further, for the embodiment, the statistical measures are employed to select **122** a distribution known to the recipient, and an identifier of the selected distribution is conveyed to the recipient (e.g. multiplexed **128** onto the transmission bit stream). In various embodiments, the distribution descriptor also serves as the identifier of the selected distribution. In

particular, it is used as an index into an array of known distribution stored at the recipient.

[0021] In various embodiments, the statistical measures determined include a mean value of the residual samples, their variances, the skewness of their distribution, and the kurtosis of their distribution. In other embodiments, the invention may be practiced with more or less statistical measures.

[0022] Still referring to Fig. 1, for the embodiment, the determined statistical measures are also employed to divide each residual sample into two portions, a most significant bits (MSB) portion, and a least significant bit (LSB) portion (to be described more fully below). For the embodiments, the LSB of each residual sample is directly transmitted to the recipient (e.g., multiplexed 128 onto the transmission bit stream without encoding). For the embodiments, the number of LSB of each residual sample being directly transmitted to the recipient is also conveyed to the recipient (e.g., multiplexed 128 onto the transmission bit stream without encoding).

[0023] In various embodiments, the mean DC offset, if applicable, is also computed, and conveyed to the recipient (e.g., multiplexed 128 onto the transmission bit stream without encoding). For these embodiments, DC offset is subtracted from the residual samples.

[0024] Further, the MSBs of each residual sample are encoded 122-124 using codewords (or simply, codes) constructed using the selected distribution. The encoded MSBs are then provided to the recipient (e.g., multiplexed 128 onto the transmission bit stream without encoding). In various embodiments, the constructed codes may be Huffman codes, run length codes, adaptive arithmetic codes, non-adaptive arithmetic codes (e.g. Gilbert-Moore codes), or other codes of the like.

[0025] In the foregoing description, the conveyance to the recipient (e.g., multiplexed 128 onto the transmission bit stream) of the various values, block sizes, prediction order, quantization sizes, quantized prediction coefficients, distribution identifier, distribution descriptor, the number of LSB of each residual sample to be conveyed, the LSB, the encoded MSB, and so forth, are immediately described following the description of their generations. The order of presentation is merely for ease of understanding. The order of these descriptions is not to be read as limiting on the

invention, requiring their conveyance on generation. The generated values may be stored, and processed into a transmission bit stream in batch. Further, multiple transmission bit stream, and/or multiple channels (of like or different kinds) may be employed for the transmission.

[0026] Referring now to the statistical measure determination **118**, the LSB identification **126**, and the MSB encoding **124** operations again, an embodiment of these operations will be described in further detail. Recall the received audio data are first partitioned **104** in to blocks or sub-blocks, and the blocks/sub-blocks are selected for processing, one block/sub-block at time. Assume the selected unit (block/sub-block) has a size of n, and the residual samples of this unit are $e_1...e_n$.

[0027] For the earlier described embodiment, where four statistical measures, the mean value of the residual sample, their variances, the skewness of the their distribution, and the kurtosis of their distribution, are computed, the computations are performed in accordance with the following formulas:

[0028] mean value of the residuals: $\bar{e} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$;

[0029] standard deviation of the residual's distribution: $\sigma = \sqrt{\mathrm{var}\,e}$, where

$$\mathrm{var}\,e = \frac{1}{n-1}\sum_{i=1}^{n}(e_i - \bar{e})^2 \;;$$

[0030] skeweness of the distribution: $\mathrm{skew}\,e = \dfrac{1}{n}\sum_{i=1}^{n}\left[\dfrac{e_i - \bar{e}}{\sigma}\right]^3$; and

[0031] kurtosis of the distribution: $\mathrm{kurt}\,e = \dfrac{1}{n}\sum_{i=1}^{n}\left[\dfrac{e_i - \bar{e}}{\sigma}\right]^4 - 3$.

[0032] Further, the distribution descriptor is formed as follows (with the quantized versions) of these quantities:

[0033] $\mathrm{dsc} = \mathrm{dsc}\left(\bar{e}, \log_2 \sigma, \mathrm{skew}\,e, \mathrm{kurt}\,e\right)$.

[0034] In alternate embodiments, e.g. embodiments offering low-complexity modes, $\mathrm{kurt}\,e = 0$; $\mathrm{skew}\,e = 0$; $\bar{e} = 0$ instead of calculating them properly.

[0035] Further, in various alternate embodiments, parameter σ may be estimated by using absolute deviation or absolute mean of the residual:

[0036] (a) $\sigma' = C_1 \dfrac{1}{n} \sum\limits_{i=1}^{n} \left| e_i - \overline{e} \right|$ or.

[0037] (b) $\sigma'' = C_1 \dfrac{1}{n} \sum\limits_{i=1}^{n} \left| e_i \right|$ (under assumption that $\overline{e} \to 0$),

[0038] where C_1 is a constant chosen in view of the distribution (e.g. for zero mean Laplacian distribution, C_1 may be set to 1/sqrt(2)).

[0039] Further, in various embodiments, the distribution descriptor may be formed using

[0040] (i) only variance estimate (e.g. when mean = 0) (optionally, using e.g. the (b) variance estimate approach described above),

[0041] (ii) variance + mean estimates (optionally, using e.g. the (c) variance estimate approach described above).

[0042] In various embodiments, on determination of the statistical measures, and selection of the distribution, an inverse-quantized mean value $\langle \overline{e} \rangle$, and the logarithm of standard deviation $\log_2 \sigma$ of the distribution $\langle \log_2 \sigma \rangle$ are reconstructed.

[0043] Then, the reconstructed values are employed to split each residual sample into MSB and LSB as follows:

[0044]

$$e_i^{MSB} = \left(e_i - \langle \overline{e} \rangle \right) >> \max\left(<\log_2 \sigma> - C_2, 0 \right);$$

$$e_i^{LSB} = \left(e_i - \langle \overline{e} \rangle \right) \& \left(\left(1 << \max\left(<\log_2 \sigma> - C_2, 0 \right) \right) - 1 \right);$$

[0045] where C_2 is an empirically selected constant. In various embodiments, C_2 is set to equal 3.

[0046] During decoding, each residual sample will be recombined as follows:

[0047] $e_i = e_i^{MSB} << \max <\log_2 \sigma> - C_2, 0)) + e_i^{LSB} + \langle \overline{e} \rangle$.

[0048] where C_2 is an empirically selected constant. In various embodiments, C_2 is set to equal 3.

[0049] In various embodiments, the distribution descriptor is also the identifier of the selected distribution, as it indexes into an array of pre-stored distributions of the MSBs of the residual samples, at both the sender and the recipient.

[0050] In various embodiments, the ranges of MSBs in these pre-stored distributions are restricted to $[-21,21]$, which approximately corresponds to the range of $[-3\sigma, 3\sigma]$ in the non-normalized distribution.

[0051] As described earlier, the selected distribution is used to construct block codes for encoding of the MSBs of the residual samples. Their LSBs will be transmitted directly using $\max\left(\langle \log_2 \sigma \rangle\ C_2, 0\right)$ bits for each residual sample. To encode samples which MSBs fall outside the $[\ -\overline{e}\ \{MSB\}_{max},\ \overline{e}\ \{MSB\}_{max}]$ range, for the embodiment, the encoder transmits an escape code $(\overline{e}\ \{MSB\}_{max} + 1)$, and then uses any standard monotonic code (e.g. Golomb codes, Golomb-Rice codes, Levenstein code, etc.) to transmit the difference $\left|e_i^{MSB}\right|$ and the escape code.

[0052] Referring now to **Figure 2**, wherein a system having a transmit section and a receive section, both adapted to practice the compression method of **Figure 1**, in accordance with some embodiments, is shown. In alternate embodiments, a system may comprise only the transmit section or the receive section. It is not necessary to always practice the invention with both sections.

[0053] As illustrated, for the embodiments, system **200** comprise transmit section **202** including transmitter **216**, and a receive section **222** including receiver **226**. In alternate embodiments, transmit and receive sections may share a common transceiver. Further, for the embodiments, in addition to transmitter **216**, transmit section **202** includes controller **218**, whereas in addition to receiver **226**, receive section **222** includes controller **230**, to control the operations of the various elements of the respective sections. Similarly, in alternate embodiments, transmit and receive sections may share a common controller instead.

[0054] For the embodiment, in addition to transmitter **216** and controller **218**, transmit section **202** further includes first selector **206**, filter **208**, encoder **212**, computer unit **210**, and second selector **214**, coupled to each other, and to transmitter **216** and controller **218** as shown. Selector **206** is employed, under the control of controller **218** to partition a portion of a stream of audio signal data into blocks or sub-blocks. Filter **208** is a prediction filter, to be applied, under the control of controller **218**, to the current of audio signal data to be processed and compressed. Compute unit **210** under the

control of controller **218** is employed to perform the various earlier described computations. Encoder **212** is employed under the control of controller **218** to encode the MSB of the residual samples as earlier described. Second selector **214** under the control of controller **218** is employed to select the various output values to be outputted, and multiplexed them onto the transmission bit stream.

[0055] For the embodiments, in addition to receiver **226** and controller **230**, receive section **222** further includes decoder **228**, and recombiner **232**, coupled to each other, and to receiver **226** and controller **230** as shown. Decoder **228**, under the control of controller **230**, is employed to decode the encoded MSB of the residual samples as earlier described. Recombiner **232** is employed, under the control of controller **230**, to recombine the received MSB and LSB to reconstitute the residual sample.

[0056] Except for the logic provided to these elements and/or their usage to cooperate with other elements to effectuate the desired compression of audio signal, these elements otherwise may be implemented in a variety of manners, in hardware, firmware, software, or combination thereof. Thus, system **200** represents a broad of range of systems having audio transmission and/or audio reception capabilities. For examples, system **200** may be a wireless mobile phone, a palm-sized computer, a tablet computer, a laptop computer, a desktop computer, a server, a set-top box, an audio/video entertainment unit, a music player, a DVD player, a CD player, a CAMCORDER, and so forth.

[0057] Thus, a novel audio signal data compression method and apparatus has been described. Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that a wide variety of alternate and/or equivalent implementations may be substituted for the specific embodiments shown and described, without departing from the scope of the present invention. This application is intended to cover any adaptations or variations of the embodiments discussed herein. Therefore, it is manifestly intended that this invention be limited only by the claims and the equivalents thereof.